Leadership role brief

VP, Research & Engineering

Lead the engineering systems and research infrastructure behind rigorous frontier AI evaluations.

About Solumn AI

Solumn AI works to understand and evaluate how frontier AI models behave in challenging, real-world environments. We develop sophisticated evaluation environments, research infrastructure, and rigorous methodologies to assess AI systems against frontier labs' safety and alignment standards.

Our goal is to translate AI safety research into practical assets—evaluations and training data—that help drive measurable improvements in model behavior.

The opportunity

We are hiring a hands-on technical leader to own our engineering direction and help shape what we research next. This is not a management-only VP role. You will build core systems, make consequential architectural decisions, lead exceptional engineers, and work with the founders to turn AI safety questions into rigorous experiments.

What you will lead

01 **Engineering platform and operations**

Architect scalable infrastructure for environment provisioning, model execution, experiment orchestration, data pipelines, observability, evaluation quality, and reproducibility.

02 **Frontier evaluation environments**

Lead the development of realistic agentic environments: tool-use and multi-agent systems, software and cybersecurity settings, long-horizon tasks, sandboxed execution, and robust verifiers.

03 **Research engineering**

Co-develop new directions in AI safety and alignment. Turn hypotheses into evaluation designs and controlled experiments; scrutinize metrics, diagnose failure modes, and contribute to credible external outputs.

04 **The technical team**

Hire and mentor engineers and research engineers. Set high standards for architecture, testing, technical review, and execution while remaining hands-on with difficult problems.

What you bring

Engineering fundamentals	Strong computer science and systems engineering judgment, demonstrated through technically complex systems shipped into production.
Zero-to-one ownership	Experience building from scratch, ideally in an early-stage startup, and evidence of making a meaningful system work at scale.
Depth in modern AI	Practical understanding of LLMs, transformers, inference, agents, tool use, and evaluation—well beyond API integration.
Research judgment	Ability to read research, question experimental assumptions, identify confounders, and turn uncertain ideas into rigorous prototypes.
Hands-on leadership	Technical ownership, mentorship, and sound decision-making. A prior VP title, large-team management experience, or academic publications are not prerequisites.

The first 180 days

First 60 days	Own the technical landscape, identify critical architectural risks and take responsibility for a high-priority environment or platform capability.
First 120 days	Deliver reusable evaluation infrastructure, stronger verification and QA, and a clear technical operating model for the team.
First 180 days	Enable Solumn to independently create and scale novel evaluation environments, and co-lead at least one substantial research initiative with an externally credible output.

Why Solumn

The role combines genuine architectural ownership, demanding frontier-AI problems, and the opportunity to help shape the research direction and technical foundations of a growing AI safety company.